Title: Audio Content Analysis in The Presence of Overlapped Classes

- A Non-Exclusive Segmentation Approach to Mitigate Information Losses

Duraid Y. Mohammed, Philip J. Duncan, Muhammad M. Al-Maathidi, Francis F. Li
d.y.mohammed@edu.salford.ac.uk , P.J.Duncan@salford.ac.uk ,
M.M.Abd@edu.salford.ac.uk , f.f.li@salford.ac.uk

School of Computing, Science and Engineering
University of Salford
Salford, Greater Manchester, UK

Soundtracks of multimedia files are information rich, from which much content-related metadata can be extracted. There is a pressing demand for automated classification, identification and information mining of audio content. A segment of the audio soundtrack can be either speech, music, event sounds or a combination of them.There exist many individual algorithms for the recognition and analysis of speech, music or event sounds, allowing for embedded information to be retrieved in a semantic fashion. A systematic review shows that a universal system that is optimised to extract the maximum amount of information for further text mining and inference does not exist. Mainstream algorithms typically work with a single class of sound, e.g. speech, music or even sounds and classification methods are predominantly exclusive (detects one class at a time) and losing much of information when two or three classes are overlapped.

A universal open architecture for audio content and scene analysis has been proposed by the authors. To mitigate information losses in overlapped content, non-exclusive segmentation approaches were adopted. This paper is presented from one possible implementation deploying the universal open architecture as a paradigm to show how the universal open architecture can integrate existing methods and workflow but maximise extractable semantic information.
In the current work, overlapped content is identified and segmented from carefully tailored feature spaces and a family of decision trees are used to generate a content score. Results show that the developed system, when compared with well established audio content analysers, can identify and thus extract information from much more speech and music segments. The full paper will discuss the methods, detail the results and illustrate how the system works.

Biography

Duraid Y. Mohammed is a 2nd year Ph.D student at university of Salford/ School of Computing, Science and Engineering/ Acoustics research centre. He has published 1 conference paper in Audio Engineering Society (AES) 136th and his second paper is accepted for published in IEEE journal.